DOCUMENT RESUME

ED 480 062 CG 032 635

AUTHOR	Perlman, Carole L.
TITLE	Practice Tests and Study Guides: Do They Help? Are They Ethical? What Is Ethical Test Preparation Practice?
PUB DATE	2003-08-00
NOTE	12p.; In: Measuring Up: Assessment Issues for Teachers, Counselors, and Administrators; see CG 032 608.
PUB TYPE	Information Analyses (070)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	*Educational Testing; *Ethics; *Study Guides; Study Skills; *Test Coaching; Test Results; *Test Wiseness
IDENTIFIERS	*Practice Tests

ABSTRACT

The increasing use of high-stakes tests has focused attention on test preparation activities. The term test preparation can apply to a number of different practices that vary on the degree to which they are defensible. Although some test preparation is legitimate, there are concerns that certain test preparation activities may have a negative impact on students' education by causing narrowing of the curriculum and overemphasis on test-taking skills and particular assessment formats. This chapter explores these topics and provides suggesting for promoting good test preparation practice. (Contains 19 references.) (GCP)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 480 062

Practice Tests and Study Guides: Do They Help? Are They Ethical? What is Ethical Test Preparation Practice?

By
Carole L. Perlman

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

CG032635

BEST COPY AVAILABLE

Chapter 27
Practice Tests and Study Guides
Do They Help? Are They Ethical? What Is Ethical Test Preparation Practice?

Carole L. Perlman

The last decade has seen increasing reliance on standards-based instruction and on assessments that measure students' mastery of instructional standards. Currently, nearly every U.S. state requires at least some K–12 students to participate in state assessments and 27 states use state assessment results for accountability purposes; some state assessments are used to determine which students may be promoted or awarded a diploma (Olson, 2001). The No Child Left Behind Act will require annual reading and mathematics testing for students in grades three through eight, with serious consequences for schools whose students score poorly. Because the stakes can be high for both students and school staff, schools may place considerable emphasis on activities designed to help students perform well on tests. However, as the American Educational Research Association points out in its Position Statement Concerning High-Stakes Testing in PreK–12 Education (2000), high-stakes testing can result in inappropriate methods of test preparation.

What Is Test Preparation?

Within the context of elementary and high school achievement testing, *test preparation* has no single, universally agreed-on definition but instead refers to a number of practices that vary in the degree of specificity with which they address a particular test. Some of the less specific forms of test preparation include teaching general strategies for taking different types of tests, teaching content from the domain being tested, and practicing with items in various formats that measure the domain tested. More test-specific strategies include practice with items in a similar format to those on the test, using state- or district-provided sample items, practice with commercial test preparation materials, practice with parallel forms or old tests, or even practice with items from the actual test—even though that is clearly

inappropriate, unethical (Cizek, 2001), and in some places, illegal (Johnston, 1999; Texas Education Code, Chap. 39, Subchap. C, ß101.65; Florida Statutes Title XVI, Chap. 228, ß228.301; the relevant sections of these state statutes may be downloaded from www.tea.state.tx.us/rules/tac/chapter101/ch101c.html and www.flsenate.gov/Statutes/index.cfm?mode=View%20Statutes&SubMenu=1&App_mode=Display_Statute&Search_String=&URL=CH0228/SEC301.HTM).

Does Test Preparation Help?

Some test preparation activities do appear to be beneficial. Mehrens and Kaminski (1989) cite several meta-analyses of test preparation research and conclude that test-wiseness training can improve scores on achievement tests, though not necessarily to a great extent. For example, practice on items similar to those on the real test can be helpful. Research also suggests that familiarizing students with the answer sheet format; encouraging them to do their best, to skip difficult items, and to listen carefully to the test directions; and giving them strategies for dealing with test anxiety *are* helpful. There is little documentation of efficacy for commercial test preparation materials for K–12 achievement tests. Interestingly enough, some commercial test preparation materials offer conflicting advice. For example, on reading tests, some encourage students to read the reading selection before looking at the questions. Others encourage students to read the questions first, which in some instances means that they are encouraging the students *not* to follow the test publisher's directions. This inconsistency can be confusing to students who have been directed to pay careful attention to the test directions. Research also suggests that reading the questions first may actually lower some students' performance (Bishop, 1999; Perlman, Borger, Gonzalez, & Junker 1988).

Those contemplating test preparation programs should remember that there are opportunity costs associated with them. Time spent on test preparation activities often comes at the expense of instruction in the content areas being assessed. The money spent on those materials might also be used in other, more productive, ways.

Effects of Test Preparation on Test Validity

The purpose of achievement testing is to make accurate inferences about what students know and can do with respect to a broad content

domain from which the test items constitute a sample of all the questions that could be asked. These inferences may form the basis for evaluating instruction, promoting students, granting merit bonuses to school staff, allocating financial and human resources, providing remedial assistance, or placing students in other special programs. Curiously, test preparation can both improve validity and decrease it.

To the extent that they increase scores without increasing the underlying subject-area knowledge and skills, test preparation activities compromise test validity. Such activities as practice with alternate forms of the test may artificially inflate students' scores, perhaps so much that a student might not receive needed remedial assistance. By providing instruction and practice only on items that mimic actual test items, we risk students not being able to generalize to the broader content area. Shepard (2000) discusses controlled studies suggesting that students who can answer a particular question correctly might not be able to answer the question if it is phrased in a slightly different way. For example, imagine that students have seen subtraction items only in this format:

Subtract: 832
<u>– 459</u>

We cannot necessarily conclude that a student who responded to this item correctly would perform equally well on the problem when presented in these formats:

$832 - 459 = ?$ or Solve for n: $832 - 459 = n$

Certainly, we would want to know whether our students can solve a problem of this nature regardless of how it is presented.

In contrast, learning about the test format and reducing anxiety might improve the validity of the scores (Messick, 1982, cited in Heubert & Hauser, 1999). The National Research Council recommends that "students should receive sufficient preparation for the specific test so their performance will not be adversely affected by unfamiliarity with its format or by ignorance of appropriate test-taking strategies" (Heubert & Hauser, 1999, p. 7).

Ethical Considerations in Test Preparation Practices

It seems clear that some test preparation practices are more defensible than others, but there is less than complete agreement among educators on which are and are not appropriate. Certainly many teachers

and principals would not agree with Mehrens and Kaminski's (1989) contention that whereas it is appropriate to briefly teach some general test-taking skills and to give general instruction on all district or state standards, it is unethical to provide practice with parallel forms of the test, to restrict instruction to the content measured by the test, or to assess students only with items similar in format to those on the test. Popham (1991) takes a similar position. School districts routinely encourage school staff to teach what the test measures, however, and many states provide extensive practice tests in order to familiarize students and teachers with test content and format. Is it wrong to use them? What is our responsibility to students' immediate and long-term interests, particularly on high-stakes tests? What kind of behavior should we model for our students? What are our professional responsibilities?

Several publications aim to clarify what practices are ethical and appropriate for professionals involved in testing: the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1999), the *Code of Professional Responsibilities in Educational Measurement* (NCME, 1995), the *Code of Fair Testing Practices in Education* (JCTP, 2002), and the *Rights and Responsibilities of Test Takers* (JCTP, 1998). The purpose of these publications is to foster fair and valid measurement that enables school staff and others to draw correct conclusions about what students know and can do. Thus, these statements provide a basis for evaluating what test preparation practices are appropriate.

Why is it so important to draw the right conclusions? Drawing the wrong conclusions might result in a student not receiving needed help or it might place a student in a course or program for which he or she is not prepared and is unlikely to succeed. It might result in the failure to allocate human and material resources to places where they are most needed. It might result in schools mandating use of inappropriate or ineffective instructional programs. It might provide false information about the course of education reforms and may lead to ill-advised policy decisions.

Some testing opponents might contend that high-stakes testing is so inherently injurious that any attempt to help students get higher scores is acceptable. Popham (1992) rejects that argument as specious, stating that regardless of the quality of the test or the way scores are used, "educators still are responsible for providing test preparation that is both professionally ethical and educationally defensible" (p. 17). Both Cizek (2001) and Popham point out that cheating by school staff sets a bad example for students and conveys the message that cheating is acceptable.

All four professional guidelines specify that test takers should be informed of the purpose of the test and given general information about the content and format of the test. The *Standards for Educational and Psychological Testing* and the *Code of Fair Testing Practices* state that all test takers should be informed of any test preparation materials that are available, and of test-taking strategies that might be either beneficial or detrimental. According to the *Standards*,

> Test-taking strategies, such as guessing, skipping time-consuming items, or initially skipping and then returning to difficult items as time allows, can influence test scores positively or negatively. Differential use of such strategies by test takers can affect the validity and reliability of test score interpretations. . . . The use of such strategies by all test takers should be encouraged if their effect facilitates performance and discouraged if their effect interferes with performance. (Standard 11.13, p. 116)

An effort should be made to make test preparation materials equally available to all examinees. If calculators, computers, or other equipment is used in testing, test takers should have the opportunity to familiarize themselves with that equipment, unless such practice would compromise the validity of the tests.

The *Standards* and the *Code of Professional Responsibilities in Educational Measurement* direct educators to refrain from engaging in test preparation practices that would lead to invalid scores. The *Code of Professional Responsibilities* enjoins test developers from marketing test preparation materials that "may cause individuals to receive scores that misrepresent their actual levels of attainment." It is the responsibility of those who select tests to "avoid recommending, purchasing, or using test preparation products and services that may cause individuals to receive scores that misrepresent their actual levels of attainment." Those who administer assessments should "avoid actions or conditions that would permit or encourage individuals or groups to receive scores that misrepresent their actual levels of attainment." The language in the *Standards* is similar: "The integrity of test results should be maintained by eliminating practices designed to raise test scores without improving performance on the construct or domain measured by the test." The authors comment that "such practices may include teaching test items in advance, modifying test administration procedures, and discouraging or excluding certain test takers from taking the test. These practices can lead to spuriously high scores that do not reflect performance on

the underlying construct or domain of interest" (Standard 15.9, p. 168). The *Standards* further stipulate that "test users have the responsibility to protect the security of tests, to the extent that the test developers enjoin users to do so" (Standard 11.7, p. 115) and the "responsibility to respect test copyrights" (Standard 11.8, p. 115). These standards clearly preclude school staff from duplicating secure test materials and from divulging test items or answers to students.

Finding a Middle Ground

It is necessary to draw a distinction between "teaching the test" and "teaching *to* the test." The former, which is never acceptable, implies disclosing actual test questions ahead of time or providing answers to questions that will appear on the actual test. The latter involves teaching the student the broad content that the test is intended to measure and may include some training in test-taking skills. A number of authors (e.g., Borger et al., 1996; Kilian, 1992; Mehrens & Kaminski, 1989; Perlman, 2000; Popham, 1992; Miyasaka, 2000, cited in Vaughn, 2001) recommend that most test preparation be integrated as seamlessly as possible into regular classroom instruction, rather than becoming a time-consuming add-on that takes the place of instruction in the content being assessed. This is consistent with the advice offered in the National Research Council report *High Stakes: Testing for Tracking, Promotion, and Graduation* (1998):

> The preparation of students plays a key role in appropriate test use. It is not proper to expose students ahead of time to items that will actually be used on their test or to give students the answers to those questions. Test results may also be invalidated by teaching so narrowly to the objectives of a particular test that scores are raised without actually improving the broader set of academic skills that the test is intended to measure. The desirability of "teaching to the test" is affected by test design. For example, it is entirely appropriate to prepare students by covering all the objectives of a test that represents the full range of the intended curriculum. We therefore recommend that test users respect the distinction between genuine remedial education and teaching narrowly to the specific content of a test. At the same time, all students should receive sufficient preparation for the specific test so their performance will not be adversely affected by unfamiliarity

with its format or by ignorance of appropriate test-taking strategies. (pp. 6–7)

This statement suggests that use of sample items is not only acceptable but desirable. Fairness dictates that the sample items be made available to *all* students, especially when the stakes are high. Teachers and students should know what kinds of assessments will be used, and sample items can be an efficient way to communicate that to every student who will be taking the test. In *High Stakes*, the National Research Council recommends that "test users should balance efforts to prepare students for a particular test form against the possibility that excessively narrow preparation will invalidate test outcomes" (p. 280).

Promoting Good Practice

In preparing students to take tests, I recommend following these guidelines:

- Provide all students with the opportunity to learn the subject area to be tested.
- To the extent possible, integrate test preparation with regular classroom instruction throughout the year.
- Assess each student's thinking skills on a daily basis. The majority of standardized test items require students to apply critical thinking skills. The more accustomed students are to doing that, the easier it will be for them to do well on the test. Borger and colleagues (1996) provide suggestions for easy ways to create homework assignments, discussion questions, and classroom assessments that require students to exercise thinking skills. In addition, students should be asked to explain how they arrived at their answers.
- In the classroom, use a variety of assessment formats rather than only the one that appears on the test. Answering open-ended questions in class can be useful preparation for taking multiple-choice tests.
- Allow students to become familiar with the test format and mechanics of test taking, but avoid spending much time on test-taking skills. Often all that is needed is brief practice with the test mechanics. Although it is desirable to familiarize students with the test format, an overemphasis on test-taking strategies may be detrimental in that it reduces the amount of time available for meaningful instruction in

the content areas to be assessed. The best test preparation is solid instruction aimed at increasing students' knowledge of the subject being tested. No amount of instruction in test-taking skills is likely to provide enough help for a student who lacks knowledge of the subject being tested.

- Avoid devoting class time to extensive review of material students have already learned.
- Discuss with students the importance of doing their best on tests.
- Provide explicit written guidelines and training on what constitutes appropriate and inappropriate practices for preparing students for tests and administering tests.
- Select appropriate tests and avoid putting too much weight on any single test.

As Vaughn (2001) points out, "instruction targeted at increasing student content mastery is not only the most ethical approach, but also addresses the overall goal of improving student achievement. . . . The most ethical and appropriate approach to test preparation is, in fact sound instructional practice" (p. 4).

Summary and Conclusion

The increasing use of high-stakes tests has focused attention on test preparation activities. The term *test preparation* can apply to a number of different practices that vary in the degree to which they are defensible. Test preparation activities may either increase test validity or reduce it. Although some test preparation is legitimate, there are concerns that certain test preparation activities may have a negative impact on students' education by causing narrowing of the curriculum and overemphasis on test-taking skills and particular assessment formats. Test preparation is best integrated into regular classroom instruction. Appropriate test preparation can include brief practice of test-taking skills and familiarization with the test format, but much greater emphasis should be placed on teaching students the curriculum standards and thinking skills that the assessments are intended to measure.

References

AERA. (2000). AERA position statement concerning high-stakes testing in preK–12 education. Retrieved from www.aera.net/about/policy/stakes.htm.

AERA, APA, & NCME. (1999). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

Bishop, N. S. (1999). The effects of different test-taking conditions on reading comprehension test performance. Paper presented at the annual meeting of the National Council on Measurement in Education, Montreal.

Borger, J. B., Cano, M., Collins, C. B., Evans, W. J., Perlman, C. L., Qualls, J. W., & Wood, J. A. (1996). *Preparing your elementary students to take standardized tests.* Chicago: Chicago Public Schools. Also available from http://intranet.cps.k12.il.us/Assessments/Preparation/preparation.html.

Cizek, G. J. (2001). Cheating to the test. *Education Next.* Retrieved from www.educationnext.org/2001sp/40.html.

Heubert, J. P., & Hauser, R. (Eds.). (1999). *High stakes: Testing for tracking, promotion, and graduation.* (Report of the Committee on Appropriate Test Use, Board on Testing and Assessment, National Research Council). Washington D.C.: National Academy Press.

JCTP. (1998). *Rights and responsibilities of test takers: Guidelines and expectations.* Washington, DC: American Psychological Association.

✦JCTP. (2002). *Code of fair testing practices in education.* Available on *Measuring Up: An Anthology of Assessment Resources* [CD]. Also retrievable on-line: http://aac.ncat.edu.

Johnston, R. C. (1999, March 17). Texas presses districts in alleged test-tampering cases. Retrieved from www.edweek.org/ew/ew_printstory.cfm?slug=27texas.h18.

Kilian, L. (1992). A school district perspective on appropriate test-preparation practices: A reaction to Popham's proposals. *Educational Measurement: Issues and Practices, 11*(4), 13–15.

Mehrens, W. A., & Kaminski, J. (1989). *Using commercial test preparation materials for improving standardized test scores: Fruitful, fruitless or fraudulent?* Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans.

✦ NCME. (1995). *Code of professional responsibilities in educational measurement.* Washington, DC: Author. Available from www.natd.org/Code_of_Professional_Responsibilities.html.

Olson, L. (2001). Finding the right mix. *Quality Counts 2001.* Bethesda, MD: *Education Week.*

Perlman, C. L. (2000). *Surreptitious inclusion of good teaching in test preparation activities.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Perlman, C. L., Borger, J., Gonzalez, C., & Junker, L. (1988). *Should they read the questions first? A comparison of two test-taking strategies for elementary students.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Popham, W. J. (1991). Appropriateness of teachers' test-preparation practices, *Educational Measurement: Issues and Practice, 10*(4), 12–15.

Popham, W. J. (1992). The perils of responsibility sharing. *Educational Measurement: Issues and Practices, 11*(4), 16–17.

Shepard, L. (2000, Winter). Why is "teaching the test" a bad thing? *State Education Leader, 18*(1), 6–7.

Vaughn, E. S. (2001, Spring). Ethical and appropriate test preparation. *ERS Spectrum,* 1–6. Retrieved from www.ers.org/spectrum/.

✦ Document is included in the Anthology of Assessment Resources CD

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

[ ] This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

[X] This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").